10/575710
IAP15 Rec'd PCT/PTO 13 APR 2006
PCT/FI2004/000611

Method and Equipment for Performing Aggregate-portion-specific Flow Shaping in Packet-Switched Telecommunications

The present invention relates to a method, according to Claim 1, for performing aggregate-portion-specific flow shaping in packet-switched telecommunications.

The invention also relates to equipment, according to Claim 3, for performing aggregate-portion-specific flow shaping formation in packet-switched telecommunications.

In this publication, the following abbreviations are used in the depiction of both the prior art and the invention:

CBS	Committed Burst Size (greatest permitted burst size [bit]), when exceeding the committed information flow,
CIR	Committed Information Rate (greatest permitted mean traffic speed [bit/s]),
FIFO	First In, First Out discipline,
MP	Measuring Point, in which the speed characteristics of the traffic flow are measured (e.g., mean traffic speed, momentary traffic speed, bust size),
PIR	Peak Information Rate (greatest permitted momentary traffic speed [bit/s]),
PKS	Packet size in bits,
V1, V2,...	Traffic flow 1, 2,...,
VTS	The earliest moment in time, after which the next packet representing a specific traffic flow, or shaping group may be forwarded, in order that not even one rule set for a speed property of the traffic flow or shaping group in question will be broken (Valid Time to Send),
VTSi	The earliest moment in time, after which the next packet representing the traffic flow i may be forwarded, in order that not even one rule set for a speed property of the flow in question will be broken,
VTSk	The earliest moment in time, after which the next packet representing the shaping group k may be forwarded, in order that not even one rule set for a speed property of the shaping group in question will be broken,

VTS_pk	The earliest moment in time, after which a packet may be forwarded, in order that not even one rule set for a speed property dealing with any shaping group associated with the packet will be broken (Valid Time to Send),
VTS_CIR	The earliest moment in time, after which the next packet representing a specific traffic flow, or shaping group may be forwarded, in order that the greatest permitted mean speed and/or the greatest permitted burst size will not be exceeded,
VTS_PIR	The earliest moment in time, after which the next packet representing a specific traffic flow, or shaping group may be forwarded, in order that the greatest permitted momentary speed will not be exceeded.

In packet-switched telecommunications systems, it is often advantageous for it to be possible to monitor and limit the speed properties of the traffic flow formed of packets being transmitted. A speed property can refer, for example, to the mean traffic speed (CIR), the size of the burst (CBS) by which the mean speed can be momentarily exceeded, or the momentary speed (PIR). The traffic flow can consist of, for example, packets to be routed to a specific transfer link, packets sent by a specific end-user, which have a specific source address, or packets to be routed to a specific transfer link, which have a specific service class. In the examinations presented in this publication, an individual traffic flow consists of packets, which are directed to a specific queue, in the order of their entry to the system, Figures 1, 2, 3, and 4.

Figure 1 shows one way, according to the prior art, of monitoring and limiting the speed properties of a traffic flow. In the following examination that further elucidates the matter, the term speed properties is used to include mean speed (CIR [bit/s]), burst size (CBS [bit]), by which the mean speed can be exceeded temporarily, and momentary speed (PIR [bit/s]). When packets are begun to be transferred past the measurement point (MP) marked in Figure 1 (i.e. when the first bit of the packet being transferred passes the point MP), the values of the variables VTS_CIR and VTS_PIR are calculated for the next packet. VTS_CIR gives the earliest permitted moment in time, when the next packet can be begun to be transferred past MP, in order not to exceed the CIR or the CBS restrictions. Correspondingly, VTS_PIR gives the earliest permitted moment in

time, when the next packet can be begun to be transferred past MP, in order not to exceed the PIR restriction. The equations 1 and 2 show the principle generally used for calculating the VST_CIR and the VTS_PIR values.

$$VTS\_CIR_{next} = \max\,(t\text{-}CBS/CIR,\ VTS\_CIR_{prev}) + PKS/CIR, \qquad (1)$$

and

$$VRS\_PIR_{next} = \max\,(t,\ VTS\_PIR_{prev}) + PKS_{prev}/PIR, \qquad (2)$$

in which t is the time, PKS is the size of the packet in bits, the sub-index 'next' refers to the next packet and the sub-index 'prev' prefers the packet the first bit of which is transferred past point MP at the moment t.

When all the speed properties are taken into account, the earliest permitted moment (VTS), when the next packet can begin to be transferred past point MP, is obtained from equation 3.

$$VTS = \max(VTS\_CIR,\ VTS\_PIR). \qquad (3)$$

In the rest of this publication, a system, by means of which the speed properties of the traffic flow can be monitored and limited, will be referred to as a 'shaper' (SH) while the operation, in which the speed properties of the traffic flow are monitored and limited, will be referred to as 'shaping'. An essential part of shaping is a buffer memory, in which it is possible to store the packets, which, due to speed restrictions, cannot be forwarded (i.e. past point MP) immediately after they have arrived in the system. If it is wished to ensure that shaping does not alter the transfer sequence of the packets, FIFO (first in - first out) queue discipline is applied in the buffer memory, as in Figure 1.

Figure 2 shows a system according to the prior art for implementing the shaping of several parallel traffic flows V1, ..., V5, in a situation, in which the flows in question are alternated (multiplexed) to form a single aggregate flow Va. Multiplexing can take place by applying, for example, the SFQ (Start-time Fair Queuing [1]) method. In the system shown in the figure, the speed properties of a portion formed of packets representing individual traffic flows (V1, ..., or V5) in an aggregate flow Va can be monitored and

limited. Measurement points (MP1, ..., MP5) relating to the multiplexed traffic flows (V1, ..., V5) are all at the same location in the output of the multiplexer. The measurement point MPi, only monitors packets representing the multiplexed traffic flow Vi (i = 1, ..., 5). Similarly, when calculating the earliest permitted transfer moment VTSi, only packets representing the traffic flow Vi are taken into account. The earliest permitted transfer moment VTSi can be calculated for a multiplexed traffic flow Vi, for example, as shown in equations 1, 2, and 3.

In the following, a situation is examined, in which it is desired to perform shaping aggregate-portion-specifically, for example, to permit the monitoring and limiting of the speed properties of a portion formed of packets representing the traffic flows V1 and V2 in the aggregate flow, the monitoring and limiting of the speed properties of a portion of packets representing the traffic flows V3 and V4 in the aggregate flow, and, in addition, the monitoring and limiting of the speed properties of the entire aggregate flow. In this publication, shaping of this kind is referred to as aggregate-portion-specific shaping. In the rest of the publication, such a situation will be expressed in such a way that traffic flows V1 and V2 belong to a specific shaping group, V3 and V4 belong to a specific second shaping group, and traffic flows V1, V2, V3, V4, and V5 belong to a specific third shaping group. The shaping group including traffic flows V1 and V2 and the shaping group including traffic flows V3 and V4 are included in the shaping group including traffic groups V1, V2, V3, V4, and V5. This means that shaping can be hierarchal.

In the system implemented according to Figure 2, aggregate-portion-specific shaping is performed, but with the limitation that, when examined in terms of shaping, the aggregate portions always consist of traffic representing only a single incoming traffic flow V1, V2, V3, V4, or V5.

Figure 3 shows a system according to the prior art for implementing shaping corresponding to the example situation described above. The traffic flows V1 and V2 are multiplexed to form the traffic flow V1a and the shaper SH1a performs shaping resulting in the traffic flow V1b. The traffic flows V2 and V3 are multiplexed to form the traffic flow V2a and the shaper SH2a performs shaping resulting in the traffic flow V2b. The

traffic flows V1b, V2b, and V5 are multiplexed to form the traffic flow V3a while the shaper SH3a performs shaping resulting in the traffic flow V3b.

Situations often arise, in which it is wished to ensure system transmittance with a higher privilege for one specific incoming traffic flow (e.g., V1), than for another traffic flow (e.g., V2). In the system according to Figure 3 when considering the traffic flows V1 and V2, this is traditionally resolved by controlling the multiplexer Mux1, in such a way that a packet can only be transferred to the FIFO1a queue if the queue is empty. Such an operation is referred to by the established term 'back pressure'. This prevents a situation arising, in which the queue FIFO1a could be filled up with packets representing the traffic flow V2, thus forcing the forwarding of packets representing the traffic flow V1, and which arrive later, to wait. Such operations complicate the system shown in Figure 3.

The following problems relate to the system according to Figure 3:

Problem 1) Even though the traffic flows V1b and V2b are monitored and limited with regard to the desired speed properties (e.g., CIR, PIR, CBS), the multiplexing with the traffic flow V5, performed in the multiplexer Mux3, results in there being no guarantee that the portions of the traffic flow V3b, formed of packets representing the traffic flows V1b and V2b, will be within the desired limits in terms of their speed properties. In other words, aggregate-portion-specific shaping cannot be performed.

Problem 2) The logical topology of the multiplexing and shaping system depends on how the traffic flows arriving in the system are bundled into different shaping groups. The system shown in Figure 3 is only responsible for bundling individual examples of the traffic flows into different shaping groups. In other words, the topology is arbitrary. This hampers the implementation of the shaping system. Particularly the circuit implementation (e.g., using an ASIC (Application Specific Integrated Circuit) microcircuit) becomes difficult, while the software implementation is also challenging.

The present invention is intended to eliminate the defects of the prior art disclosed above and for this purpose create an entirely new type of method and equipment for performing

shaping in packet-switched telecommunications. The aim of the invention is a method and equipment for performing shaping in such a way as to avoid the aforementioned problems relating to the prior art.

The method according to the invention is characterized by what is stated in the characterizing portion of Claim 1.

The equipment according to the invention is, in turn, characterized by what is stated in the characterizing portion of Claim 3.

In the following, the invention is examined in greater detail with the aid of examples according to the accompanying figures.
Figure 1 shows a block diagram of one way according to the prior art of monitoring and limiting the speed properties of a traffic flow.

Figure 2 shows a block diagram of a system according to the prior art for implementing the monitoring and limiting of the speed properties of several parallel traffic flows V1, ..., V5, in a situation in which the flows in question are alternated (multiplexed) to form a single aggregate flow Va.

Figure 3 shows a block diagram of a system according to the prior art concerning an example of a situation, in which it is wished to perform shaping, in such a way that it is possible to monitor and limit the speed properties in an aggregate flow of a portion formed of packets representing the traffic flows V1 and V2, it is possible to monitor and limit the speed properties in an aggregate flow of a portion formed on packets representing the traffic flows V3 and V4, and, in addition, it is possible to monitor and limit the speed properties of the aggregate flow V3.

Figure 4 shows a block diagram of a system according to the invention for performing shaping, in such a way that the traffic flows arriving at the system can belong to shaping groups in an arbitrary manner and the speed properties of the aggregate portion formed of packets representing an arbitrary shaping group can be monitored and limited (aggregate-portion-specific shaping).

The theoretical basis of the method according to the invention will become apparent from the following examination.

Traditionally, shaping is applied to a specific traffic flow, in which the packets move in a temporal sequence, for example, according to Figures 1, 2, or 3. In the method according to the invention, the central concept is a shaping group, i.e. an aggregate portion. As stated above, the shaping group consists of the incoming traffic flows, the speed properties of the portion of the aggregate flow formed of packets representing which are monitored and limited using a restriction totality of specific speed properties (such as a totality formed of the CIR, PIR, and CBS values). In the method according to the invention, the VTS value is shaping-group-specific. The VTS value of a specific shaping group thus expresses the earliest permitted moment in time, at which a packet representing the relevant shaping group can be forwarded (first bit past the measurement point), in order that a rule set for even one speed property (e.g., CIR) of the relevant shaping {group} will not be broken. The measurement point too is logically shaping-group-specific. In Figure 4, the measurement points of all the shaping groups k are at the same location. Naturally it is possible to implement shaping equipment, in which the measurement points MP of one or more shaping groups k are located separately. The following examination is restricted to a situation, in which the measurement points MP of all of the shaping groups are at the same location. It is then possible to refer simply to the measurement point, instead of to the measurement point relating to a specific shaping group.

An individual traffic flow can belong to one, several, or no shaping group k. If the traffic flow does not belong to even one shaping group k, the shaping equipment naturally will not set any limit to the speed properties of the traffic flow in question.

The earliest permitted moment, at which a single packet can be forwarded (first bit past the measurement point), is determined on the basis of all the shaping groups k, under which the traffic flow represented by the packet in question belongs. Stated more precisely,

$$VTS\_pk = \max(VTSk,\ \text{packet belongs under shaping group k}), \qquad (4)$$

in which VTS_pk is the earliest moment in time, after which the packet can be forwarded (first bit past the measurement point), in order that not even one rule set for the speed properties associated with any shaping group concerning the packet will be broken, and VTSk is the VTS value of the shaping group k.

In the system according to Figure 4, the packet being examined is the packet closest to the multiplexer (i.e. at the head of the queue) of some FIFO queue. When the packet is offered to the multiplexer 10, and if the multiplexer 10 selects the packet in question, the packet is then moved immediately past the measurement point. Thus, there is no internal transfer delay in the multiplexer equipment. Thus, VTS_pk states the earliest moment in time, at which the packet is permitted to be offered to the multiplexer. For this reason, in Figure 4, the permit/refuse operations of the transfer of the packet are set between the FIFO queues I-L and the multiplexer 10.

When the packet is begun to be moved past the measurement point, the VTS values of all the shaping groups, under which the packet in question belongs, are updated. If the restrictions of the speed properties of the shaping group are expressed in the form CIR, PIR, CBS, the VTS values can be updated, for example, as shown in equations 1, 2, and 3.

In brief, the principle of the shaping system according to the invention is as follows:

Situation:	Packet (first bit) passes measurement point.
Operations:	Update all VTS values of the shaping groups, under which the packet in question belongs.
Situation:	Determine earliest permitted moment, at which the packet can pass the measurement point.
Operations:	Seek maximum of the VTS values of the shaping groups, under which the packet in question belongs.

As can be surprisingly noticed from the above examination, the method according to the invention places no restrictions on how the traffic flows entering the system can be

grouped into different shaping groups. By locating the measurement points of all the shaping groups according to Figure 4, it is possible to monitor and limit the speed properties of an aggregate portion formed on packets representing an arbitrary shaping group.

References:

[1] Pawan Goyal, Harric M. Vin, Haichen Cheng. *Start-time Fair Queuing: A Scheduling Algorithm for Integrated Services Packet Switching Networks.* Technical Report TR-96-02, Department of Computer Sciences, University of Texas Austin.